MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1551 August, 1995

A Network Charge-oriented MOS Transistor Model

Jacob Katzenelson and Aharon Unikovski

This publication can be retrieved by anonymous ftp to publications.ai.mit.edu.

Abstract

The MOS transistor physical model as described in [3] is presented here as a *network model.* The goal is to obtain an accurate model, suitable for simulation, free from certain problems reported in the literature [13], and conceptually as simple as possible. To achieve this goal the original model had to be extended and modified. The paper presents the derivation of the network model from physical equations, including the corrections which are required for simulation and which compensate for simplifications introduced in the original physical model.

Our intrinsic MOS model consists of three nonlinear voltage-controlled capacitors and a dependent current source. The charges of the capacitors and the current of the current source are functions of the voltages V_{gs}, V_{bs}, and V_{ds}. The complete model consists of the intrinsic model plus the parasitics. The apparent simplicity of the model is a result of hiding information in the characteristics of the nonlinear components.

The resulted network model has been checked by simulation and analysis. It is shown that the network model is suitable for simulation: It is defined for any value of the voltages; the functions involved are continuous and satisfy Lipschitz conditions with no jumps at region boundaries; Derivatives have been computed symbolically and are available for use by the Newton-Raphson method. The model's functions can be measured from the terminals. It is also shown that small channel effects can be included in the model. Higher frequency effects can be modeled by using a network consisting of several sections of the basic lumped model.

Future plans include a detailed comparison of the network model with models such as SPICE level 3 and a comparison of the multi-section higher frequency model with experiments.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for this research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-92-J-4097 and by the National Science Foundation under grant number MIP-9001651.

1 Introduction

Our interest in modeling MOS transistors started from our work on the Supercomputer Toolkit [1, 2]. The Supercomputer Toolkit is a family of hardware modules (processors, memory, interconnect, and input-output devices) and a collection of software modules (compilers, simulators, scientific libraries, and high-level front ends) from which high-performance, special-purpose computing systems can be easily configured and programmed. The main advantage of the Supercomputer Toolkit is that for numerically intensive computations it provides supercomputer class performance at a fraction of a supercomputer cost. One such application is the simulation and optimization of high speed circuits with respect to physical parameters.

For optimization of Bi-CMOS circuits we needed a model of the MOS transistor. Our colleagues (see acknowledgment) recommended a long list of references, including [3, 4, 5, 6]. Tsividis [3] is the most comprehensive book of its kind. The simplest approach would have been to use the SPICE level 3 model [7]. However, several shortcomings of this model have been reported [13, 12], all of them involving the simulation of analog circuits of the type we intended to use. With the Toolkit in hand, computation time did not seem the hindering factor for the problems we intend to attach, so we decided to construct the most accurate model possible. Thus, the above references, in particular [3], became the basis for this work.

The type of transistor model we wanted was a *network model*, i.e., one that models the transistor as a network of more primitive devices. The resulting network can be use as a component, i.e., interconnected with other components to form larger networks. The appropriate transistor model depends on the nature of the required solution. In particular, for transient simulation the model, once embedded in a network satisfying certain conditions, should lead to an appropriate set of ordinary differential equations which can be integrated numerically to yield a solution (see, for example, [10]).

The material we found in the above references was not quite in the form we wanted. There was no large signal network model and only the small signal model could be described as a network model. The number of details was very large. The emphasis was on deriving the model from physical considerations and we missed a structural, hierarchical description of the results where the details, if needed, could be looked up later and filled in. The

physical equations were derived as if the transistor had infinite dimensions. The lumped models derived from these equations are approximations to the behavior of a finite size transistor. Yet, we could not find a mention of what properties could or could not be measured from the terminals. Thus, our goal was to start from physical considerations (using the above references) and find a "simple" large signal network model whose components can be measured from the terminals.

In the next section we derive the model for the usual operating region of the MOS transistor. The result is presented in figure 2. Section 3 gives our reasons for extending the model to other operating regions and then presents such extension. Section 4 presents the results graphically. Section 5 addresses the issue of measuring the characteristics externally. Section 6 discusses some properties of the model, in particular those related to the existence and uniqueness of the response of networks constructed with our transistor models as components. The last section presents the summary and the conclusions.

Appendix A presents the model equations in the "usual" operating region. Appendix B discusses the derivation of the incremental capacitance from the charge functions using symbolic computing techniques.

2 Region I – The usual operating region

This section presents the model in the usual operating region. The notations follows [3]. The treatment is brief and justification can be found in the above reference. Figure 1 presents the transistor including the definitions of the voltages and currents at its terminals.

Modeling starts by considering the behavior of the transistor in several regions defined by the terminal voltages V_{gs}, V_{bs}, and V_{ds}. Region I is defined by the following relations among the voltages:

$$V_{ds} \geq 0$$

$$V_{bs} \leq 0.$$

This region is, in turn, divided into four subregions:

Accumulation: $V_{gs} \leq V_{fb} + V_{bs}$, where V_{fb} is the flat band voltage;

Cutoff: $V_{fb} + V_{bs} < V_{gs} \leq V_{th}$, where V_{th} is the cut–off voltage;

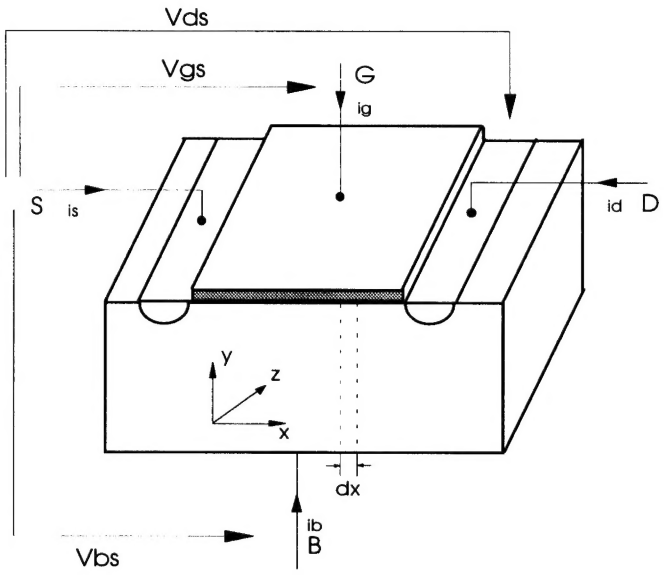

Figure 1: The MOS transistor, it's terminal voltages and currents, and a differential cross section.

Saturation: $V_{gs} > V_{th}$ and $V_{ds} \geq V_{dsat}$, where V_{dsat} is the saturation voltage;

Linear: $V_{gs} > V_{th}$ and $V_{ds} < V_{dsat}$.

In the above, V_{th} is the threshold voltage, a function of V_{bs},

$$V_{th} = V_{to} + \gamma * (\sqrt{\phi_s - V_{bs}} - \sqrt{\phi_s}),$$

and V_{dsat}, the saturation voltage, is

$$V_{dsat} = \frac{V_{gs} - V_{th}}{1 + \delta}$$

$$\delta = \frac{\gamma}{2\sqrt{1 + \phi s - V_{bs}}},$$

where V_{to}, γ and ϕ_s are physical constants (i.e., independent of the voltages) which are defined in Appendix A.

To compute the current and the charges one considers a differential subsection of the transistor (see Figure 1). Assuming the fields in the transistor are the same as if the transistor were infinite in the x and y direction, from physical consideration, one calculates the current through the subsection and the charges in it as functions of certain physical parameters and of the field intensity in the subsection. This field intensity can be expressed in terms of the voltages V_{gs}, V_{bs}, and V_{ds}.

The current, I_{ds}, is the crrent flowing through the differential cross section, in the (minus) x direction due to both drift and diffusion. It is assumed that only a negligible amount of this current enters the bulk and, therefore, the current entering the cross-section on the left exits on the right. From that also follows that its value can be obtained by integration along the the subsection face (integration along the z axis and the y axis; the distribution along the y axis being uniform as a result of the infinite dimension assumption). The integration results in the following expression for the linear and saturation regions:

$$I_{ds} = \frac{W}{2L}\mu C_{ox} V_{dsat}(V_{gs} - V_{th})(1 - \alpha^2)(1 + \lambda V_{ds}) \tag{1}$$

where W and L are the width and length of the transistor, C_{ox} is the gate oxide capacitance per unit area, α is zero in the saturation region and equal to

$$1 - \frac{V_{ds}}{V_{dsat}}$$

in the linear region. The last term, $(1 + \lambda V_{ds})$ is (alas!) a correction term introduced experimentally (it does not follow from the physical considerations in the infinite dimension x).

As we move from saturation to cutoff the transistor passes through a medium and then a weak–inversion zone; I_{ds} goes down sharply and eventually diminishes to zero:

$$I_{ds} = I_0 \exp \frac{V_{gs} - V_{th}}{\frac{3kT}{2q}}, \tag{2}$$

where

$$I_0 = \frac{W}{L}\mu C_{ox}\frac{3}{2}(\frac{kT}{q})^2 e^{-1}.$$

To make I_{ds} a continuous function of V_{gs}, I_0 has to be added to the value of I_{ds} in the saturation and linear regions (see appendix A for the complete equation).

The charges provide a more intriguing story. In the infinitesimal section dx above, four charges (per unit length) are defined:

- q_g – the charge on the metal gate;

- q_I – the inversion layer charge;
- q_b – the charge in the bulk;
- q_0 – the interface charge, a fixed quantity independent of the fields in the section.

These charges satisfy

$$q_g + q_I + q_b + q_0 = 0.$$

Each of these charges can be expressed as a function of the position x and the voltages V_{gs}, V_{bs} and V_{ds}. The charges can also be integrated over x to yield the total gate charge, Q_g, the total inversion (layer) charge, Q_I, the total bulk charge, Q_b, and the total interface charge, Q_0. The sum of the total charges is still zero, and so is the sum of the charges' derivatives with respect to time.

The above relations for the charges have been derived under DC conditions. It is assumed (the pseudo-static assumption) that as the voltages change, but "not too fast," the charges change according to the statically derived relations. (As usual, a quantitative value for "not too fast" is not given in the literature and frequency bounds on the validity of the pseudo-static model have to be obtained experimentally.)

The next step is conceptually delicate: We would like to partition Q_I into two charges Q_d and Q_s, such that $Q_I = Q_s + Q_d$; and, to associate each charge with the corresponding terminal — Q_s with the source, Q_d with the drain, Q_g with the gate and Q_b with the bulk. Moreover, we claim that, given that the proper partition is done, the time derivative of each of these charges is (a component of) the current in the corresponding terminal, i.e. $\frac{dQ_g}{dt} = i_g$, etc. Except for the case of Q_g the support of this step is not trivial; it is given in detail, together with a proper partition, in chapter 7 of [3] and its references. One of the arguments given there is that the results agree with experiment.

Notice that since

$$Q_g + Q_I + Q_b + Q_0 = 0,$$

and $\frac{d}{dt}Q_0 = 0$,

$$\frac{d}{dt}Q_g + \frac{d}{dt}Q_s + \frac{d}{dt}Q_d + \frac{d}{dt}Q_b = 0. \tag{3}$$

In words, the above equation means that the terminal currents due to changes in the charges satisfy Kirchoff's current law.

Thus, we may model the effect of the charges and I_{ds} using three voltage-controlled non-linear capacitors[1] and a dependent current source (see Figure 2). The characteristics of C_{gs}, for example, is given by the function specifying Q_g as a function of V_{gs}, V_{bs}, and V_{ds}. This means that this capacitor is a four-terminal component, needing a reference and three terminals to specify its charge (see Figure 3. The time derivative of the charge on C_{gs}, the charge Q_g, is a current entering g and leaving through s. Thus, changes in, say, V_{ds}, may change Q_g and cause a gs current. C_{bs} and C_{ds} are similarly described – each with its own characteristics. The functions are given in Appendix A. The characteristics of the current source is described by equations 1 and 2.

The choice of the non-linear capacitor, rather than the commonly chosen incremental capacitor, is a major one. It provides a simple link with the charge equations and, at the same time, simplifies the model conceptually. The incremental capacitors are the partial derivatives of the characteristics at a given operating point.

Note that we can define a capacitor for each of the terminal charges. Any three of the four can be chosen for the model. The network, or rather Kirchoff's law, guarantees that the sum of the terminal currents is zero. From equation 3 follows that the omitted charge is equal to the minus of the sum of the other three charges (plus a constant which we are at liberty to take as Q_0). Note also that in Figure 2 the source s was chosen as a reference node and the charge, say, Q_g was associated with the capacitor C_{gs}. This choice yields characteristics whose expressions corresponds directly to the charge equations in the literature [3]. If, however, the base or gate is chosen as the reference node, symmetry implies that the characteristics of C_{sb} is equal to the one of C_{db}.

The model of Figure 2 is called the the intrinsic model. To this model we have to add the so-called "parasitics". These are two linear capacitors representing the overlap between the gate metallic layer on both the source and the drain, respectively, and two nonlinear capacitors representing the effect of the pn junctions between the source to bulk and drain to bulk. A

[1] A capacitor is called *voltage-controlled* if its charge is given as a function of its voltage, i.e., $q = f(v)$; The function is called *the characteristics* of the capacitor. In our case, v is a vector whose components are voltages. The extension of the terminology is straight forward.

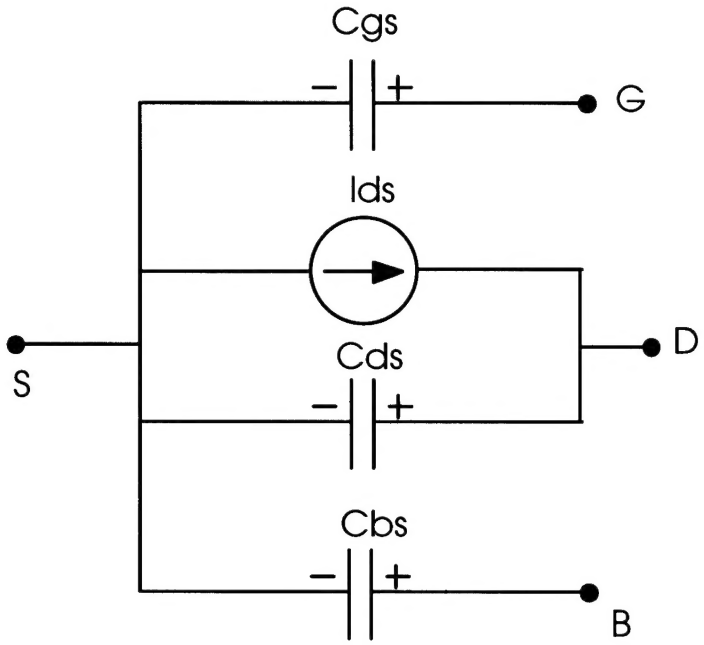

Figure 2: The intrinsic model

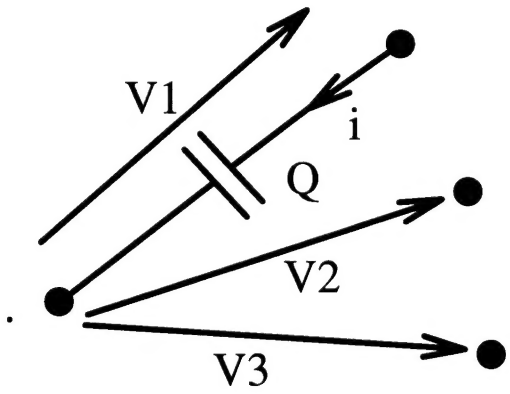

Figure 3: A four-terminal capacitor; Q is a function of (V_1, V_2, V_3); The current $i = \frac{d}{dt}Q$.

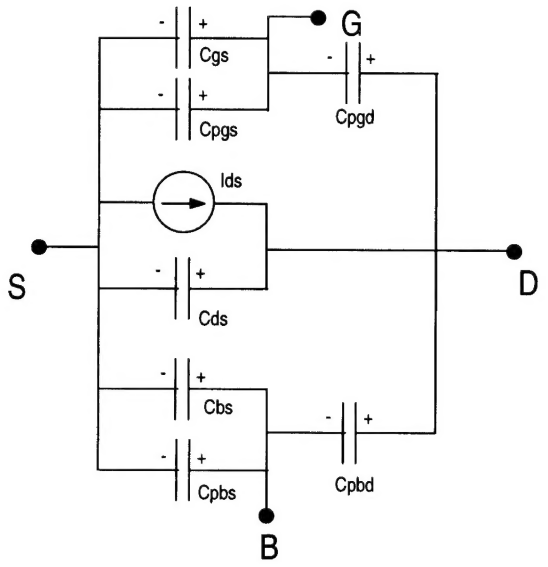

Figure 4: The complete model; The intrinsic model plus parasitics.

better model for the *pn* junction is that of a non-linear capacitor in parallel with a non-linear resistor which represents a diode; In region I, however, the diode is reversed biased, so the resistor can be neglected.

Figure 4 describes the complete model. The functions describing the characteristics of the components appear in Appendix A.

Discussion: The above contains an outline of the modeling process. Most of the details are tucked away in the functions that describe the component characteristics. We would like to consider these functions from the point of view of the model use, i.e., circuit simulation. We expected the functions to satisfy Lipschitz conditions in terms of the independent variables [2] which, here, are the voltages. Extracting the functions from literature we came across several surprises. The literature provided algebraic expressions for each region (i.e., the saturation region, cutoff, etc.) but the functions were not continuous at the boundaries. This we discovered by displaying cuts of the functions at various values. As a result, the functions in Appendix A contain corrections that we introduced heuristically.

In circuit simulation, the Newton-Raphson method is used to solve the algebraic equations that result from implicit integration methods. Newton-Raphson requires the partial derivatives of the functions with respect to the voltages; E.g., $\frac{\partial Q_g(V_{gs},V_{bs},V_{ds})}{\partial V_{gs}}$. Deriving these twelve functions was not simple. The expressions in each region are complicated enough so we just could not

[2] In order to have a solution defined by a set of first order differential equations the network has to satisfy certain conditions. One of these conditions is that the functions satisfy the Lipschitz conditions in term of the independent variables.

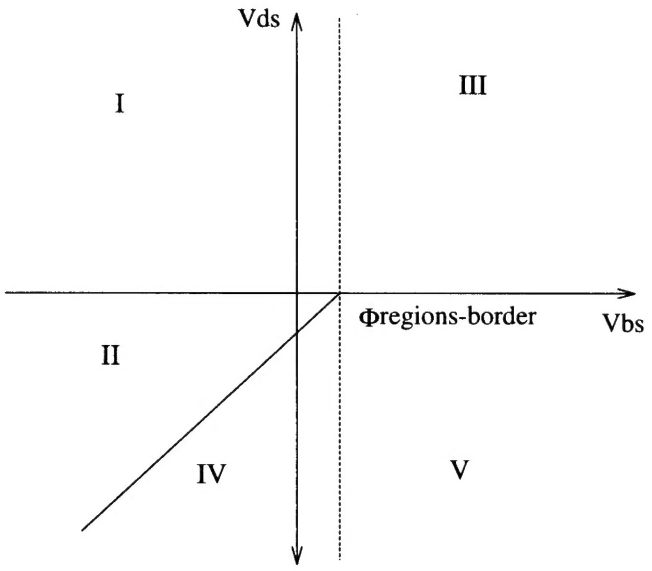

Figure 5: The five regions. Lines (except the V_{ds} coordinate) are boundaries between regions.

get the evaluation right when we did the differentiation by hand. We found that we had a wrong function on hand by testing. Every solution, i.e., proposed derivative function, was compared with the numerically approximated value; when the discrepancy was too large it indicated a mistake.

Thus, we turned to symbolic algebra to evaluate the derivatives. The straightforward derivation, however, produced enormous expressions which we could not simplify well. We circumvented this difficulty by writing a symbolic program for evaluating each derivative. This technique is described in Appendix B by example. Figure 18 depicts the partial derivative of Q_g with respect to V_{gs}.

It is worthwhile noting that these derivatives have circuit meaning. They are the incremental capacitance and trans-capacitance that appear in the small signal model. There are nine of those — the partial derivatives of Q_g, Q_b, and Q_d each with respect to each of the voltages V_{gs}, V_{bs}, and V_{ds}.

On inspection one sees that these derivatives are not continuous. The discontinuities occur along the subregions' boundaries. In all cases, the derivative normal to the boundary has the same sign on both side of the boundary. This fact is aesthetically displeasing. It does not, however, effect the existence and uniqueness of the transient solution. In our circuit examples it did not affect the convergence of the Newton-Raphson iteration. It does, however, indicate a need to make the small signal model more accurate at

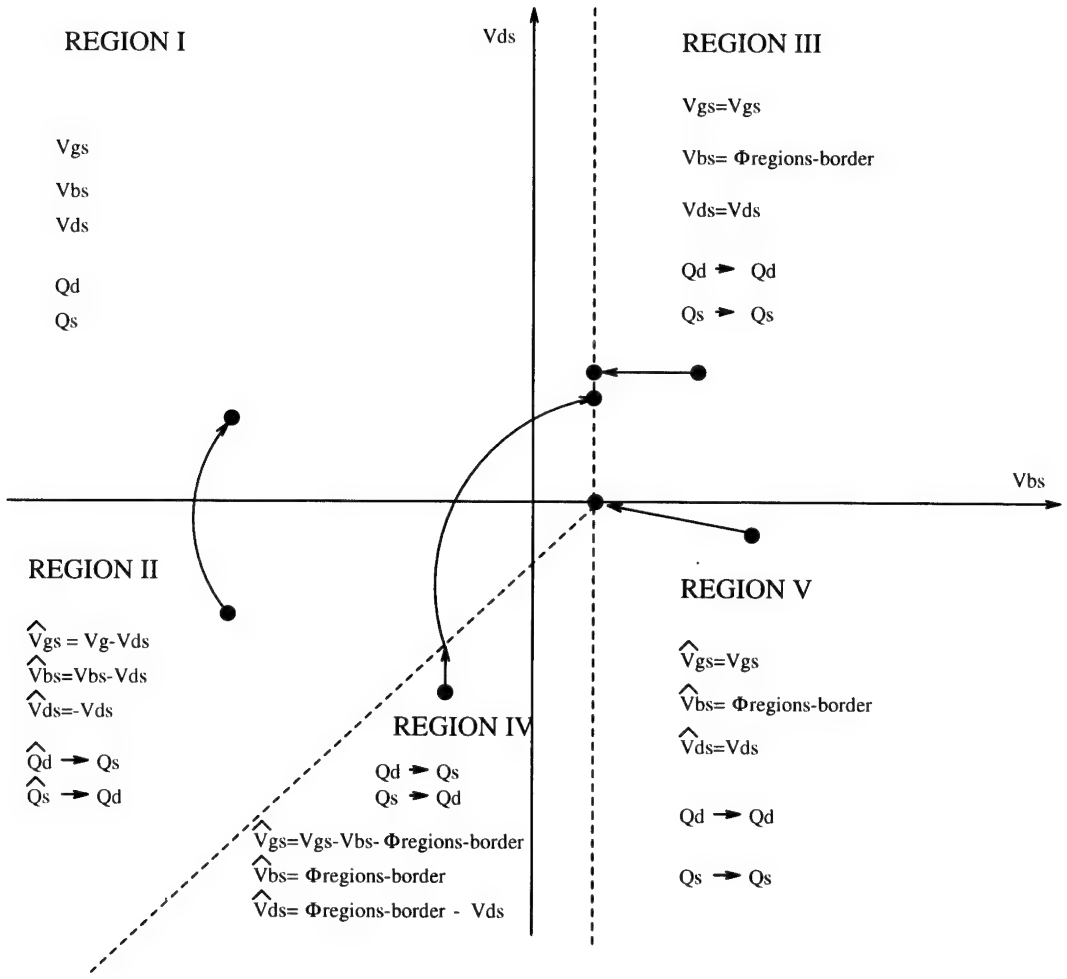

Figure 6: A picturial illustration of evaluating the functions Q_d and Q_s in each region. In region I, Q_d and Q_s are evaluated from V_{gs}, V_{bs} and V_{ds}. In region II the $\widetilde{V}$'s are evaluated first. This amounts to translating V_{gs}, V_{bs} and V_{ds} of region II to a point in region I illustrated by the curved arrow from region II to region I. Q_d and Q_s are evaluted at that region I point; The value of Q_d in region II is the value just obtained for Q_s in region I. In other regions the functions are evaluated similarly.

region boundaries.

3 The Five Regions Model

We have coded the Region I model and run some simple examples. Again, to our surprise, the simulation trajectory crept out of Region I into regions where the model was not valid. A check of a MOS circuit (An inverter with active p-channel load, [11] page 391) showed that indeed the actual circuit leaves region I. The phenomena is not a feature of the model or of the simulation program but appears in actual circuits. Therefore, we constructed a model that is valid for all values of V_{gs}, V_{bs}, and V_{ds}. We refer to this model as the *five regions model.* Figure 5 gives the five regions that partition the (V_{bs}, V_{ds}) space. For each region we define expressions that are valid for any value of V_{gs}.

3.1 Region I

We return to region I in order to extend this region somewhat. Considering the functions in region I we find that they behave properly for $V_{bs} > 0$ except that some of the derivatives become infinite (blow up) for $V_{bs} = \phi_s$. Thus, region I can be extended to

$$V_{ds} > 0$$

$$V_{bs} \leq \phi_{regions-border};$$

where $\phi_{regions-border} < \phi_s$. In our simulations we took $\phi_{regions-border} = 0.5\phi_s$.

The point $V_{ds} = 0$, $V_{gs} = V_{th}$, i.e., $V_{dsat} = 0$, is rather delicate. The various functions are continuous but the numerical evaluation is apt to "blow up" as a result of devision by zero if care is not taken. In our own representation of the function the phenomina occurs in the calculation of *alpha* as

$$\alpha = 1 - \frac{V_{ds}}{V_{dsat}} \quad for \quad V_{ds} \leq V_{dsat}$$
$$\alpha = 0 \quad for \quad V_{ds} > V_{dsat}$$

caused a division by zero for V_{ds} zero. A slight change in the conditions eliminated the numerical trouble:

$$\alpha = 1 - \frac{V_{ds}}{V_{dsat}} \qquad for \qquad (V_{ds} \leq V_{dsat}) and (V_{dsat} > 0)$$

$$\alpha = 0 \qquad otherwise.$$

The fact that this change helped depends on the numerical properties of our system, in particular, the way the system rounds a small number to zero. In general, note that at the above point functions such as Qg have a double zero at the numerator and a single zero at the denominator and the numerical result might depend on the order of computation.

3.2 Region II

The region is defined by the symmetry of the MOS transistor. Clearly, the source and drain are symmetric and for appropriate voltages the source acts as drain and the drain acts as source. Region II is the region where we would be in region I if we interchanged the source and the drain. Let us name this replacement of source by drain "the exchange operation." Denote the voltages before the exchange by V, and the voltages after the exchange by $\tilde{V}$. The relation between the the V's and the $\tilde{V}$'s can be simply concluded from Figure 1:

$$\tilde{V}_{gs} = V_{gs} - V_{ds},$$
$$\tilde{V}_{bs} = V_{bs} - V_{ds},$$
$$\tilde{V}_{ds} = -V_{ds}.$$

It is easy to verify that these equations map the following region into region I:

$$\{(V_{gs}, V_{bs}, V_{ds}) | (V_{bs} < \phi_{regions-border}) and (0 > V_{ds} \geq V_{bs} - \phi_{regions-border})\}$$

The charges and currents that we want to evaluate in region II can be computed using the corresponding charge and current functions of region I:

$$Q_g(V_{gs}, V_{bs}, V_{ds}) = Q_g(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds});$$

$$Q_b(V_{gs}, V_{bs}, V_{ds}) = Q_b(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds});$$

$$Q_d(V_{gs}, V_{bs}, V_{ds}) = Q_s(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds});$$

$$Q_s(V_{gs}, V_{bs}, V_{ds}) = Q_{dIn}(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds});$$

while

$$I_{ds}(V_{gs}, V_{bs}, V_{ds}) = -I_{ds}(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds}).$$

Special care has to be taken when derivatives are calculated. To wit:

$$\frac{\partial Q_g}{\partial V_{ds}}(V) = -\frac{\partial Q_g}{\partial V_{gs}}(\tilde{V}) - \frac{\partial Q_g}{\partial V_{bs}}(\tilde{V}) - \frac{\partial Q_g}{\partial V_{ds}}(\tilde{V}).$$

The procedures evaluating the functions were written as the above implies: Given the V's, the $\tilde{V}$ is computed and the value of the appropriate function in region I is calculated; the signs, etc., are next adjusted according to the evaluated function.

3.3 Regions III, IV, and V

We know next to nothing of the behavior of the intrinsic transistor outside regions I and II; nor could we find any literature on it. Since we needed the functions to be define over thewhole domain, we extended the definitions of the basic functions, I_{ds}, Q_s, Q_b, Q_d, and Q_g, outside regions I and II to satisfy the following conditions: the functions should be continuous and satisfy the Lipschitz conditions over the entire three dimensional space. In addition, the extension should be simple.

Figure 6 describes the the partition of the space into regions and illustrates the evaluation of Q_d and Q_s in each region. Given a point (V_{gs}, V_{bs}, V_{ds}) in region III, IV, or V, a point $(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds})$ is calculated on the boundary of regions I or region II. The value of a function, one of I_{ds}, Q_s, Q_b, Q_d, or Q_g, at (V_{gs}, V_{bs}, V_{ds}) is the value of the function at $(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds})$.

The definitions of the regions and the transformations are as follows

- Region III:

$$\{(V_{gs}, V_{bs}, V_{ds}) | (V_{bs} > \phi_{regions-border}) and (V_{ds} \geq 0)\}$$

$$\begin{aligned} \tilde{V}_{gs} &= V_{gs}, \\ \tilde{V}_{bs} &= \phi_{regions-border}, \\ \tilde{V}_{ds} &= V_{ds}. \end{aligned}$$

- Region IV:

$$\{(V_{gs}, V_{bs}, V_{ds}) | (V_{bs} < \phi_{regions-border}) and (V_{ds} < V_{bs} - \phi_{regions-border})\}$$

$$\begin{aligned} \tilde{V}_{gs} &= V_{gs}, \\ \tilde{V}_{bs} &= V_{bs}, \\ \tilde{V}_{ds} &= V_{bs} - \phi_{regions-border}. \end{aligned}$$

- Region V:

$$\{(V_{gs}, V_{bs}, V_{ds}) | (V_{bs} > \phi_{regions-border}) and (V_{ds} < 0)\}$$

$$\begin{aligned} \tilde{V}_{gs} &= V_{gs}, \\ \tilde{V}_{bs} &= \phi_{regions-border}, \\ \tilde{V}_{ds} &= 0 \end{aligned}$$

Discussion: No claim is made that the above equations capture the behavior of the intrinsic transistor in regions III, IV and V. As will be shown in the sequel, they give us a model that has a unique solution under some reasonable conditions and, thus, when the simulation enters these region we can emit a warning message and continue instead of ending with a programming bug.

4 Graphical presentation of the results

There is a basic difficulty in displaying functions of three variables as are the current and charges of the intrinsic transistor. To gain some insight into the properties of these functions we did the following: Consider figure 6 as a plane, $V_{gs} = constant$, in the space (V_{gs}, V_{bs}, V_{ds}). Chose a value for V_{bs} which defines the dotted line (the path) in the figure. Figures 10 to 18 depict the current I_{ds}, the charges, and some small signal parameters commonly

used by analog circuits designers, along this path. The region boundaries are shown as vertical lines in the figures.

The values that we have chosen and which are common to all the figures are as follows:
$V_{bs} = -5V$;
V_{ds} varies continuously from -10V to 10V;
V_{gs} is a parameter value; The values that were chosen for it are 0V, 2V, 4V, 6V and 8V.

The following parameters were chosen for the representation of the results:
$\frac{W}{L} = 1$
$d_{ox} = 700\mathring{A}$ (Oxide thickness)
$N_a = 3 \cdot 10^{15}\ cm^{-3}$ (Doping concentration)
$L_{ov} = 0.1\mu$ (Channel length overlap)
$\Phi_{ms} = -0.93V$ (Metal semiconductor potential)
$Q_0 = 0\ cm^{-3}$ (Effective oxide charge)
$\mu_{0n} = 1200\ \frac{cm^2}{v \cdot sec}$ (Electron mobility)
$\mu_{0p} = 250\ \frac{cm^2}{v \cdot sec}$ (Hole mobility)
$V_{sat} = 1 \cdot 10^5\ \frac{m}{sec}$ (Saturation velocity)
The path that was chosen shows the behavior of the functions in region I and II, the most commonly used regions, and in region IV.

5 Determining the model's characteristics by measurement

One major property that we expect models of physical devices to have is the ability to measure experimentally the device characteristics from the device's terminals, and then to use the characteristics in simulation. The SPICE level 3 model, for example has this property. In this section we present a way of measuring the model characteristics from the terminals.

Given a MOS transistor device, we would like to extract Q_g, Q_b, Q_d, Q_s and I_{ds} as functions of V_{gs}, V_{ds}, and V_{bs}. Figure 4 shows some parasitic capacitors in addition to the capacitors of the intrinsic model. These capacitors represent the capacitance of the source-bulk and drain-bulk *pn* junctions, and the overlap capacitance between the gate and the source, and between the gate and the drain.

When measuring the charge-voltage characteristics for each of the tran-

sistor's terminal pairs, it is not possible to distinguish between the charge that is supplied to the parasitic capacitor and the charge that is supplied to the intrinsic capacitor that appears in parallel to it without prior assumptions on the behavior of the voltage-charge characteristics of each one of these separately. Thus, in this paper we measure the two charges together. The measurement can be done by either small signal capacitive measurement (C-V measurement) or large signal charge voltage measurement. In C-V measurement we get the incremental capacitance and, in the sequel, it is assumed that this is the method used.

Let C_{ij} be the total incremental capacitance between nodes i and j in the the device. Thus, we get:

$$C_{gs} = C_{pgs} + \left.\frac{\partial Q_g}{\partial V_{gs}}\right|_{V_{bs},V_{ds}},$$

where C_{pgs} is the parasitic capacitance between the gate and the source (the overlapping capacitance);

$$C_{gd} = C_{pgd},$$

where C_{pgd} is the parasitic capacitance between the gate and the drain (the overlapping capacitance);

$$C_{ds} = \left.\frac{\partial Q_d}{\partial V_{ds}}\right|_{V_{gs},V_{bs}}$$

$$C_{bs} = C_{pbs} + \frac{\partial Q_b}{\partial V_{bs}}$$

C_{pbs} is the source–bulk pn junction capacitance;

$$C_{bd} = C_{pdb}$$

C_{pdb} is the drain bulk pn junction capacitance. We denote by $\widetilde{C_{ij}}$ the incremental capacity measurement between terminals i and j when the other terminals are connected to a constant DC level.

The measurement results for each set of V_{bs} V_{gs} and V_{ds} are:

$$\widetilde{C_{gs}} = C_{gs} + C_{gd}$$

$$\widetilde{C_{bs}} = C_{bs} + C_{bd}$$

$$\widetilde{C_{ds}} = C_{ds} + C_{db} + C_{gd}$$
$$\widetilde{C_{gd}} = C_{gd} + C_{db} + C_{ds}$$
$$\widetilde{C_{gb}} = C_{bs} + C_{db}$$

The matrix form of these results is:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} C_{gs} \\ C_{gd} \\ C_{ds} \\ C_{db} \\ C_{bs} \end{bmatrix} = \begin{bmatrix} \widetilde{C_{gs}} \\ \widetilde{C_{bs}} \\ \widetilde{C_{ds}} \\ \widetilde{C_{gd}} \\ \widetilde{C_{gb}} \end{bmatrix} \tag{4}$$

We would like to obtain the incremental capacitance vector vs. the measured capacitance vector, hence we get:

$$\begin{bmatrix} C_{gs} \\ C_{gd} \\ C_{ds} \\ C_{db} \\ C_{bs} \end{bmatrix} = \begin{bmatrix} 1 & -0.5 & -0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & -0.5 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -0.5 & 0.5 & -0.5 & 1 \\ 0 & 0.5 & -0.5 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{C_{gs}} \\ \widetilde{C_{bs}} \\ \widetilde{C_{ds}} \\ \widetilde{C_{gd}} \\ \widetilde{C_{gb}} \end{bmatrix} \tag{5}$$

Therefore, we can get the incremental capacitances of the model's components by measurements. In order to get the characteristics, charges vs. voltages, we need one point on the characteristics in addition to the increments. One such point is: For $V_{gs} = V_{fb}, V_{ds} = V_{bs} = 0$, $Q_g = -Q_0$, $Q_d = Q_b = Q_s = 0$.

These measurements can be used to compare the theoretical derived model with the actual behavior of the transistor, or to obtain the characteristics of the model experimentally.

6 Some properties of the of the model

This section describes some qualitative properties of the model all of which related to simulations. These properties are invertability, satisfiability of Lipschitz conditions, monotoneity of the model and the existence and uniqueness of solution of networks containing such models as components.

Is the model charge controlled? We are accustomed to expect the model to be charge–controlled [4]. The expressions defining the charges and currents in our model are all expressed in terms of the voltages V_{gs},V_{bs},and V_{ds}. Does the function which defines the charges in terms of the voltages have an inverse?

Simple inspection of the intrinsic model yields that (a) the expressions in regions I and II are too complicated for eye-ball conclusions, and (b) The functions are "flat" on the other regions and, therefore, no inverse of the intrinsic model can exist.

If, however, the parasitics are added to the picture and the model is considered as a whole, the situation is not obvious. So, we consider the model with its parasitics, i.e., the linear overlap capacitors C_{pgs} and C_{pgd}, and the pn junction capacitors $Cpbs$ and C_{pbd}. Since C_{pbs} and C_{pbd} are not define outside regions I and II they were extended by a small linear capacitors defined over the whole domain. Now, the answer is no longer obvious. The next step is the use of a global inverse theorem and an experimental verification of the conditions. Let us choose the source s as a reference node, and let Q_g, Q_b,and Q_d denote the node charges (cut-set charges associated with the nodes) which include the charge of the parasitics. Let h be the mapping from the voltages to the charges, i.e.,

$$(Q_g, Q_b, Qd) = h(V_{gs}, V_{bs}, V_{ds}).$$

The model is charge-controlled if h has a global inverse. To prove that h has the desired property we use the following theorem.

Palais's Theorem [16]: Let h be a differentiable function from R^n to R^n. Then h is a diffeomorphism [3], provided that h satisfies the two conditions

$$\begin{aligned} &(i) \quad \lim_{\|x\|\to\infty} \|h(x)\| = \infty \\ &(ii) \quad \det Dh(x) \neq 0 \ for \ all \ x, \end{aligned}$$

were $Dh(x)$ is the Jacobian of h at x.

Our functions are continuous and differentiable inside the regions but are not differentiable at the regions boundary. On one hand, this implies that if an inverse exists, it would not be differentiable; On the other hand, (ii)

[3] h has a continuous differentiable inverse over the domain

provides local homoeomorphism and, therefore, we have to prove local homoeomorphism on the boundaries some other way. Now, for each boundary, derivatives on the "right" and the "left" exist. Thus, to insure local homoeomorphism it is sufficient to show that, at each point on the boundary, the corresponding partial derivatives have the same sign. Only the normal derivatives have to be checked as continuity implies that derivatives along the boundary are equal. This check was done experimentally; the derivatives were found symbolically and evaluated numerically at various points on the boundaries.

Let V be the vector of voltages and Q – the vector of charges. As $\|V\|$ sufficiently increased, so does $\|Q\|$; While the charges of the the intrinsic capacitors do not increase for any increase in V, the charges of the linear capacitors do increase. Thus, condition (i) above is satisfied.

Condition (ii) inside the regions has been checked numerically. The partial derivatives have been found symbolically and the determinant was calculated at various points in space. The result is shown by figure 7.

Thus, strictly speaking, the above are numerical evidence that the model is charge-controlled rather than mathematic proofs. A proof of (ii) can be provided by the evaluation of the determinant in each region and showing that it does not vanish.

The charges and I_{ds} as functions of the voltages satisfy the Lipschitz conditions [4].

Since all the above functions are continuous and, on each region, they are defined as expressions, the above is determined by inspection.

The cutset charges are strictly monotonic functions of the voltages.

A function f from R^n to R^n is called *monotonic* if for every $x_1, x_2 \in R^n, x_1 \neq x_2$,

$$\langle x_1 - x_2, f(x_1) - f(x_2) \rangle \geq 0.$$

[4] A function f satisfy the Lipschitz conditions on some domain D if there exists a k, $k > 0$, such that for every x_1 and x_2 in D

$$\| f(x_1) - f(x_2) \| < k \| x_1 - x_2 \| .$$

Satisfiability of Lipschitz conditions by f implies existence and uniqueness of solution of $\frac{d}{dt}x = f(x)$. For differentiable functions, a uniformly bounded derivatives implies that the Lipschitz conditions are satisfied.

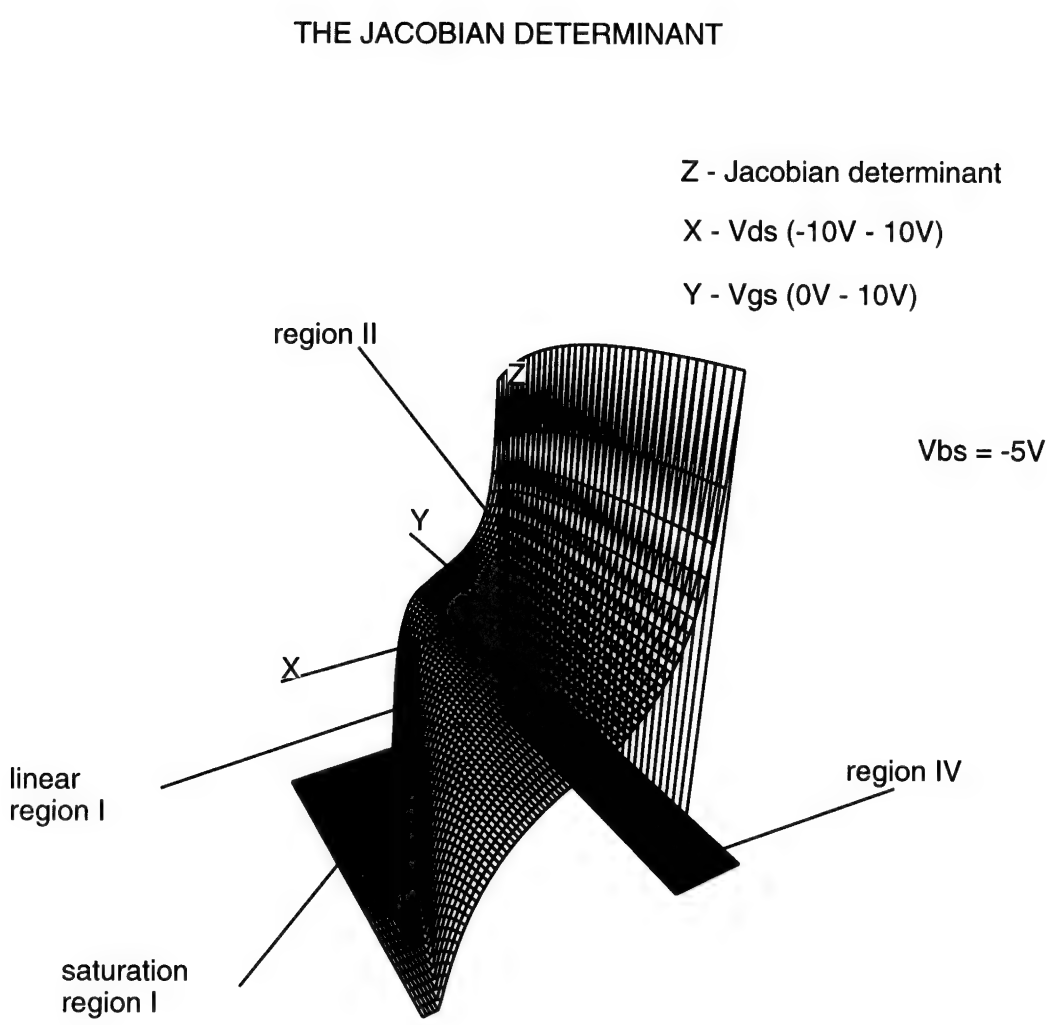

Figure 7: The jacobian determinant

In the above, if the ($\geq$) is replaced by ($>$) the function is called strictly monotonic. If the right handside is replaces by $> c\|x_1 - x_2\|^2$ the function is called strongly monotonic.

If f is one-dimensional and differentiable, the above means that the derivative of f is positive if f is strictly monotonic and positive or zero if it is monotonic.

Our interest in this property is motivated by certain fixed-point theorems on monotonic operators [17, 18]. These theorems are used in [15] to derive sufficient conditions for the existence and uniqueness of the solutions of nonlinear networks containing transistors.

The characteristics of the intrinsic transistor are monotonic but not strictly monotonic since their their derivatives in regions II IV and V are zero (see figures 8 and 9). Here the parasitic capacitors help in the following way: First, the overlap capacitors are linear and satisfy both the Lipschitz conditions and the monotonicity conditions. However, the value of the *pn* junction incremental capacitors is zero when the corresponding diode is conducting. To correct that, we added a small linear capacitor in parallel with the *pn* junction capacitor and consider the sum of these two capacitors as the parasitic capacitor. Now, if the transistor is considered as a whole, the intrinsic model plus the parasitics, it is strictly monotonic. In fact, it is even strongly monotonic with the coefficient c determined by the parasitics.

7 Summary and Conclusions

We have presented a model which is defined for any value of the terminal voltages, has continuous characteristics (no jumps) all of which satisfy Lipschitz conditions. The model is defined in terms of physical constants and the device geometry. Basically, for region I, the model summarizes the physical model which appears in [3]. Some changes were introduced, however, to make its characteristics continuous.

The model is summarized as a non-linear network with capacitors rather than with incremental capacitors as in [4, 5, 6] or, in fact, SPICE [7]. The translation of the physical model equations to capacitors is natural, although the concept requires some getting used to. The same network approach can be used to express the equations of the short channel transistor in network terms. Another extension which is conceptually simple is the modeling for

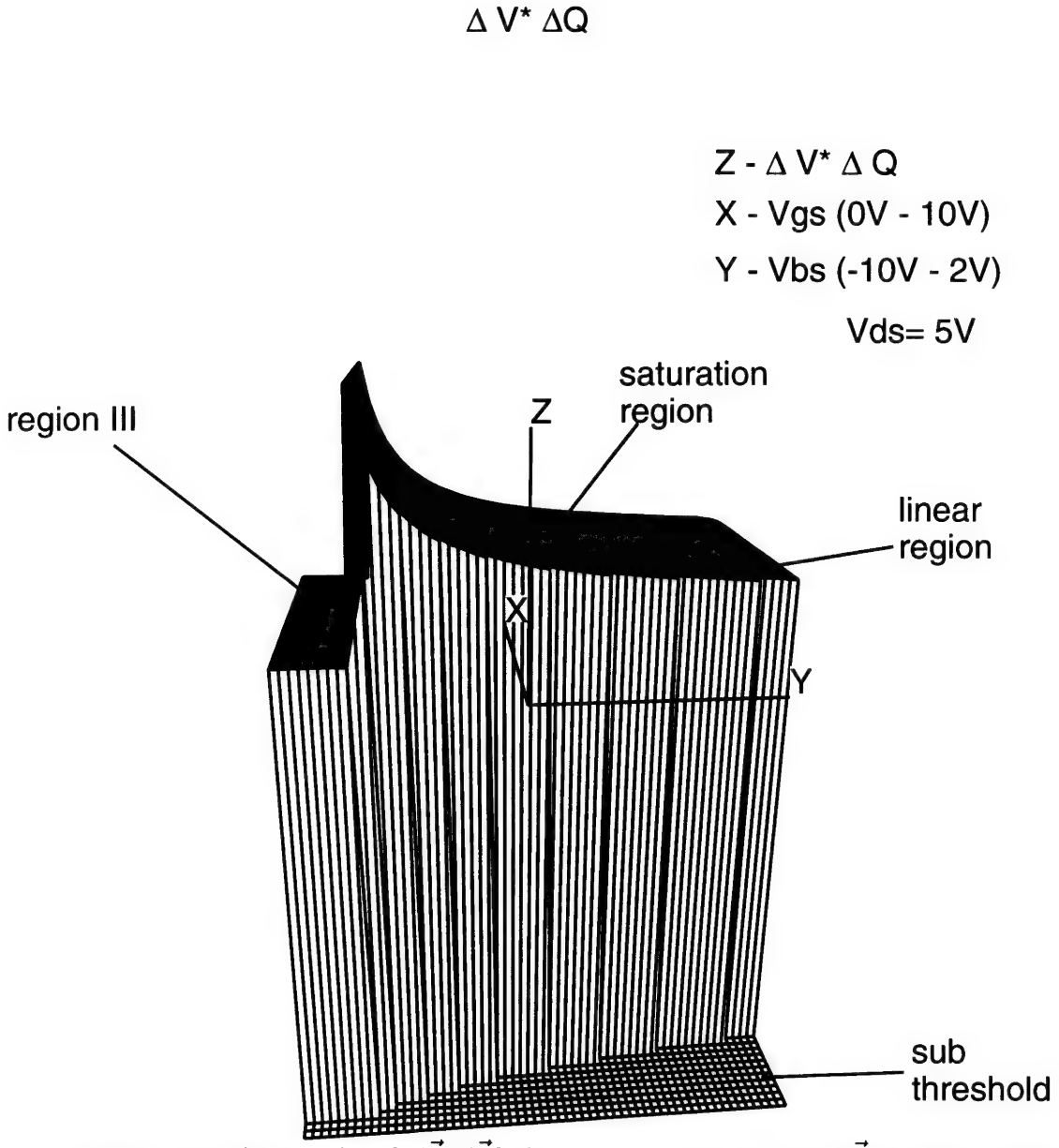

Figure 8: The result of $\Delta\vec{V}\Delta\vec{Q}$ for each component of $\Delta\vec{V}$ is $0.1V$ and $V_{ds} = 5V$

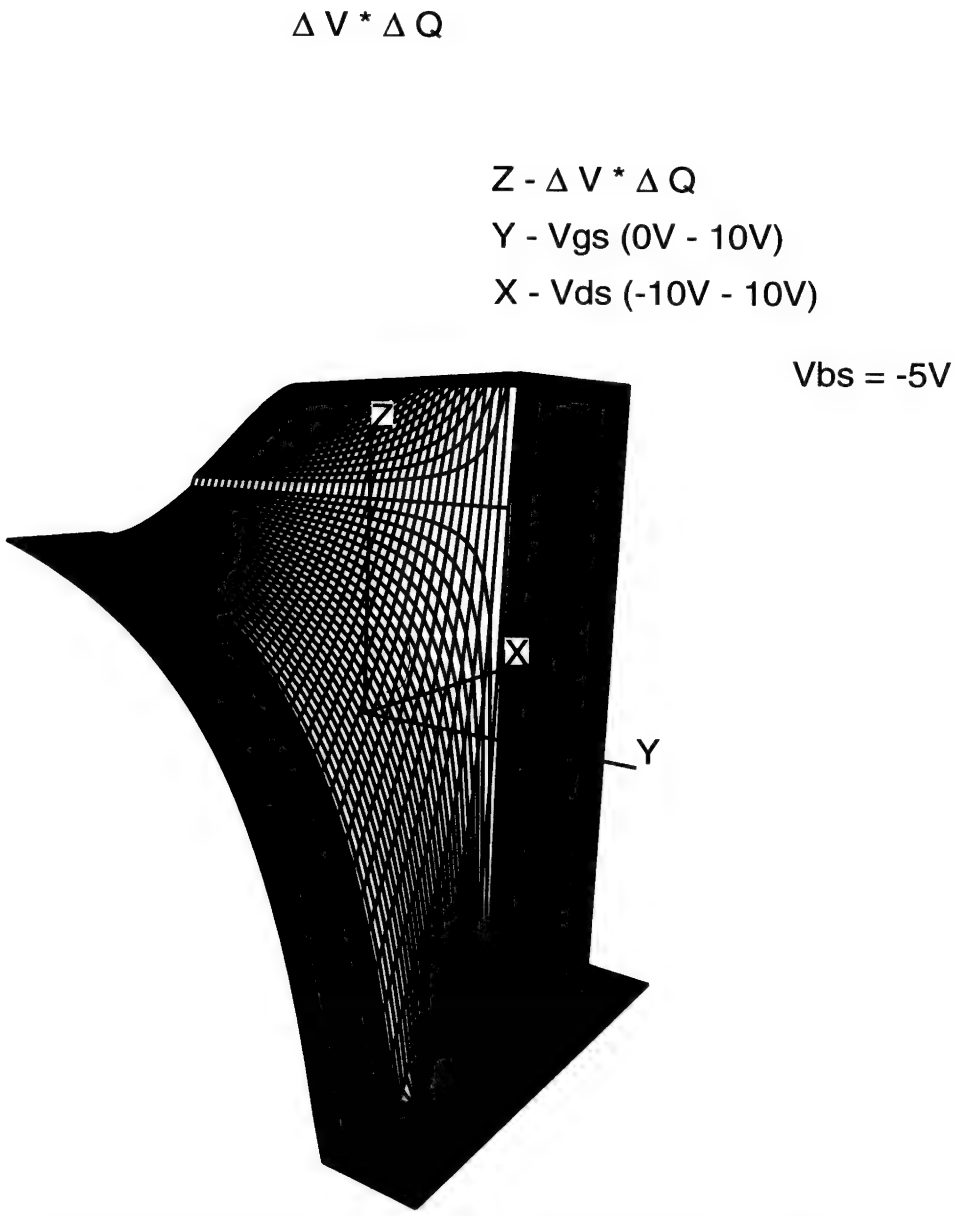

Figure 9: The result of $\Delta\vec{V}\Delta\vec{Q}$ for $\Delta\vec{V} = 0.1\vec{V}$ and $V_{bs} = -5V$

higher frequencies by dividing the transistors to cross-sections, and modeling each such section with a lumped network model similar to the network model presented here.

The model is a "large signal model" and it suffices for simulation. Moreover, conditions for existence and uniqueness of response have been developed.

By taking appropriate partial derivatives small signal behavior can be investigated around any operating point. These partial derivatives are the incremental capacitors, incremental resistors and gains. The partial derivatives are continuous everywhere but on the boundaries. I.e., as a result of continuity, the same value is obtained for the derivative 'along' the boundary when the derivative is evaluated on either side of a boundary. The normal derivative, however, evaluated on one side of the is different from the derivative evaluated at the same boundary point, but which is considered to be on the other side of a boundary. In all cases, we observed that the two derivatives have the same sign, a feature used in the proof of the existence of the inverse.

As a result of the discontinuity of the partial derivatives, the incremental capacitors, etc., are discontinuous functions of the voltages. The technique used in [8] can be used to smooth the derivatives at the boundaries.

It seems that transistors models with less jumps than in SPICE [7] are available as proprietary models in industry [9]. Since these models are private their exact properties are not clear to us. It seems that [8] and ours are the only "smooth" models publicly available.

Symbolic analysis played a major role in our work. We used it to derive the derivatives of the charges and the current and also to investigate the properties that the model satisfies.

The model parameters can be measured from the terminals. Thus, the model provides a framework for experimentally defined models. I.e., the topology and, maybe, some parameters are assumed as in the model, and the device characteristics can be measured from the terminals (see [5]). We have not conducted such measurements. This interesting work has yet to be done. Its results are expected to shade light on the model's accuracy, compare it in detail with other models, and in particular, compare the multi-section network model with a transistor operating at higher frequencies.

8 Appendix A: Expressions for currents and charges in region I

This appendix contains the expressions for the *long channel device.* Similar expressions for I_{ds} of transistors with small dimensions can be found in [3]. We did not find in the literature expressions for the charges of small dimension transistors. We have looked at the expressions for small dimensions I_{ds} and corrected the discontinuities on the boundaries. The resulting function obeys the Lipschitz conditions and can be used in the model in the same way as the long channel I_{ds}.

Some common physical relations:
Fermi level for p-type material: $\phi_F = \frac{kT}{q} \ln \frac{n_a}{n_i}$.

$$\phi_s = 2\,\phi_F$$

$$C_{ox} = \frac{\varepsilon_{ox}}{d_{ox}}$$

$$\gamma = \frac{\sqrt{2\,\varepsilon_s\,q\,N_a}}{C_{ox}}$$

$$V_{th} = V_{t0} + \gamma\left(\sqrt{\phi_s - V_{bs}} - \sqrt{\phi_s}\right)$$

where V_{t0} is:

$$V_{t0} = V_{fb} + \phi_s + \gamma\sqrt{\phi_s}$$

I_{ds} in saturation and linear regions:

$$I_{ds} = I_0 + \frac{W}{2\,L}\mu_0 C_{ox} V_{dsat}(V_{gs} - V_{th})(1 - \alpha^2)(1 + \lambda V_{ds}) \tag{6}$$

where α is defined as:

$$\alpha = 1 - \frac{V_{ds}}{V_{dsat}} \quad for \quad V_{ds} < V_{dsat}$$
$$\alpha = 0 \qquad\qquad for \quad V_{ds} \geq V_{dsat}$$

I_{ds} for the sub-threshold region, i.e., $V_{gs} < V_{th}$:

$$I_{ds} = I_0 \exp \frac{V_{gs} - V_{th}}{\frac{3kT}{2q}} \tag{7}$$

where

$$I_0 = \frac{W}{L}\mu_0 C_{ox}(\frac{kT}{q})^2 1.5e^{-1}$$

q_g for the accumulation region, $V_{gs} \leq (V_{fb} + V_{bs})$:

$$Q_g = WLC_{ox}(V_{gs} - V_{bs} - V_{fb}) - Q_0 \tag{8}$$

Q_g for $(V_{fb} + V_{bs}) < V_{gs} \leq V_{th}$:

$$Q_g = WLC_{ox}\frac{q}{qvth}\gamma\sqrt{\phi_s - V_{bs}} - Q_0 \tag{9}$$

where q and $qvth$ are:

$$q = \frac{-\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + V_{gs} - V_{bs} - V_{fb}}$$

$$qvth = \frac{-\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + V_{th} - V_{bs} - V_{fb}}$$

.

Q_g for the linear and saturation regions, $V_{gs} \geq V_{th}$:

$$Q_g = WLC_{ox}\left(V_{dsat}(\delta + \frac{2}{3}\frac{1+\alpha+\alpha^2}{1+\alpha}) + \gamma\sqrt{\phi_s - V_{bs}}\right) - Q_0 \tag{10}$$

where

$$\delta = \frac{\gamma}{2\sqrt{\phi_s - V_{bs}}}$$

$$V_{dsat} = \frac{V_{gs} - V_{th}}{1+\delta}$$

Q_b for the accumulation region, $V_{gs} \leq (V_{fb} + V_{bs})$:

$$Q_b = -WLC_{ox}(V_{gs} - V_{bs} - V_{fb}) \tag{11}$$

Q_b for $(V_{fb} + V_{bs}) < V_{gs} \leq V_{th}$:

$$Q_b = -WLC_{ox}\frac{q}{qvth}\gamma\sqrt{\phi_s - V_{bs}} \tag{12}$$

Q_b for the linear and saturation regions, $V_{gs} \geq V_{th}$:

$$Q_b = -WLC_{ox}\,[V_{dsat}\,\delta\,(1 + \frac{2}{3}\frac{1+\alpha+\alpha^2}{1+\alpha}) + \gamma\sqrt{\phi_s - V_{bs}}\ ] \qquad (13)$$

.

Q_d:
$Q_d = 0$ for $V_{gs} \leq V_{th}$.

Q_d for $V_{gs} > V_{th}$:

$$Q_d = -WLC_{ox}\,(V_{gs} - V_{th})\,\frac{4 + 8\alpha + 12\alpha^2 + 6\alpha^3}{15(1+\alpha)^3} \qquad (14)$$

Q_s:
$Q_s = 0$ for $V_{gs} \leq V_{th}$.
Q_s for $V_{gs} > V_{th}$:

$$Q_s = -WLC_{ox}\,(V_{gs} - V_{th})\,\frac{6 + 12\alpha + 8\alpha^2 + 4\alpha^3}{15(1+\alpha)^2} \qquad (15)$$

an example of a region function (for I_{ds})

- **Region I**

$$I_{ds} = f\,(V_{gs}\,,\,V_{bs}\,,\,V_{ds})$$

- **Region II**

$$\tilde{I}_{ds} = -\,f\,(\tilde{V}_{gs}\,,\,\tilde{V}_{bs}\,,\,\tilde{V}_{ds})$$

where:

$$\tilde{V}_{gs} = V_{gs} - V_{ds}$$
$$\tilde{V}_{bs} = V_{bs} - V_{ds}$$
$$\tilde{V}_{ds} = -V_{ds}$$

- **Region III**

$$\tilde{I}_{ds} = f\,(V_{gs}\,,\,\phi_{regions-border}\,,\,V_{ds})$$

- **Region IV**

$$\tilde{I}_{ds} = -f(\tilde{V}_{gs}, \tilde{V}_{bs}, \tilde{V}_{ds})$$

where:

$$\tilde{V}_{gs} = V_{gs} + \phi_{regions-border} - V_{bs}$$
$$\tilde{V}_{bs} = \phi_{regions-border}$$
$$\tilde{V}_{ds} = -\phi_{regions-border}$$

- **Region v**

$$\tilde{I}_{ds} = f(V_{gs}, \phi_{regions-border}, 0)$$

An example of a derivative function ($\frac{\partial I_{ds}}{\partial V_{gs}}$)
for $V_{gs} \leq V_{th}$

$$\frac{\partial I_{ds}}{\partial V_{gs}} = I_0 \left[\frac{V_{gs} - V_{th}}{1.5 \frac{kt}{q}} \exp(\frac{V_{gs} - V_{th}}{1.5 \frac{kt}{q}} - 1)(\frac{2/3}{\frac{kt}{q}} - \frac{2/3}{\frac{kt}{q}} \frac{\partial V_{th}}{\partial V_{gs}}) \right] \quad (16)$$

for $V_{gs} > V_{th}$

$$\frac{\partial I_{ds}}{\partial V_{gs}} = \frac{W}{2L} \mu C_{ox} (1 + \lambda V_{ds}) \cdot \quad (17)$$

$$\left\{ V_{dsat}(1 - \alpha^2) + \left(\frac{(1 - \alpha^2)(V_{gs} - V_{th})}{1 + \delta} \right) \right. \quad (18)$$

$$\left. -2\alpha V_{dsat}(V_{gs} - V_{th}) \left[\frac{V_{ds}}{V_{dsat}^2 (1 + \delta)} \right] - (1 - \alpha^2) V_{dsat} \frac{\partial V_{th}}{\partial V_{gs}} \right\} \quad (19)$$

9 Appendix C: Obtaining the Characteristics Derivative- An Example

The following example illustrates how the derivatives of the charges have been derived. The following is the scheme program specifying Q_g as a function of the voltages and the physical parameters. It is followed by a program that calculates the partial derivative of Q_g with respect to V_{gs} while all other parameters are constants. The resulting program is the third item in this appendix.

```
(define v->Qg
  (lambda ( Vgs Vbs Vds Vth mu-zero W L )
    (let ((Basic-v->Qg
           (named-lambda (Basic-v->Qg  Vgs Vbs Vds Vth mu-zero W L )

             (define (f1 vgs vbs vds)
               (+ (* W L Cox
                     (let((q (+ (- (* 0.5 gamma))
                                (sqrt (+
                                       (* gamma gamma 0.25) Vgs (- Vbs) (- Vfb)))))
                          (qvth (+ (- (* 0.5 gamma))
                                   (sqrt (+
                                          (* gamma gamma 0.25) Vth (- Vbs) (- Vfb))))))
                       (/ (* gamma (sqrt (- Phis Vbs)) q) qvth)))))

             (define (f2 vgs vbs vds)
               (let* ((delta (/ gamma (* 2. (sq-root (+ 1. (- Phis Vbs))))))
                      (Vdsat (/ (- Vgs Vth) (+ 1. delta)))
                      (alpha (if (and (<= Vds Vdsat) (> Vdsat 0.))
    (- 1. (/ Vds Vdsat));; Linear region
    0.   )));;saturation
                 (*  W L  Cox
                     (+ (* Vdsat
                           (+ delta (* 0.667 (/ (+ 1. alpha (* alpha alpha))
                                                (+ 1. alpha)))))
                        (* gamma (sq-root (- Phis Vbs)))))
                 ))

             (cond ((< Vgs (+ Vfb Vbs)) (+ (*  W L Cox (+ Vgs (- Vbs) (- Vfb)))))
                   ((< Vgs Vth) (f1 Vgs vbs vds))
                   (else (f2 vgs vbs vds)))
             )))
      (regions+ Basic-v->Qg Vgs Vbs Vds Vth mu-zero W L))))
```

The following is evaluated in the symbolic environment to yield the derivative needed. A special utility function is used: $((pd\ i)f)$ finds the partial derivative of a scheme function f with respect to the i-th variable. Let f be a lambda expression of two variables, $(((pd\ 1)f)'a\ 'b)$ finds the function with respect to the first variable and then finds the value of the derivative at the (symbolic) point a,b.

```
(define d/dvgs-Qg
  '(lambda ( Vgs Vbs Vds Vth mu-zero W L )
     (let ((Basic-d/dvgs-Qg
    (lambda (Vgs Vbs Vds Vth mu-zero W L )
```

```
       (define f1
 (lambda (vgs vbs vds)
   (* W L Cox
      (let*((q (+ (- (* 0.5 gamma))
  (sqrt (+
 (* gamma gamma 0.25) Vgs (- Vbs) (- Vfb)))))
    (qvth (+ (- (* 0.5 gamma))
     (sqrt (+
    (* gamma gamma 0.25) Vth (- Vbs) (- Vfb))))))

,(let* ((q (lambda (vgs) (+ (- (* 0.5 gamma))
    (sqrt (+
   (* gamma gamma 0.25) Vgs (- Vbs) (- Vfb))))))
(qvth 'qvth)
(exp (lambda (q)
       (/ (* gamma (sqrt (- Phis Vbs)) q) qvth))))
   (* (((pd 1) exp) 'q) (((pd 1) q) 'vgs)))))))

; cutoff 270

       (define f2
 (lambda (vgs vbs vds)
   (let* ((delta (/ gamma (* 2. (sqrt (- Phis Vbs)))))
  (Vdsat (/ (- Vgs Vth) (+ 1. delta)))
  (alpha (if (and (<= Vds Vdsat) (> Vdsat 0.))
     (- 1. (/ Vds Vdsat));; Linear region
     0.   )));;saturation

     (*  W L  Cox
 (if (> Vdsat 0.)
 ,(let* ((delta 'delta)
 (vdsat (lambda(Vgs) (/ (- Vgs Vth) (+ 1 delta))))
 (vdsat1 '(/ (- Vgs Vth) (+ 1 delta)))
 (alpha (lambda(vdsat) (- 1 (/ Vds Vdsat))))
 (exp (lambda (vdsat alpha)
(+ delta
   (* 2/3
      (/ (+ 1 alpha (* alpha alpha))
 (+ 1 alpha))))))
 (exp1
  '(+ delta
      (* 2/3
 (/ (+ 1 alpha (* alpha alpha))
```

```
    (+ 1 alpha))))))
    (+ (* exp1 (((pd 1) vdsat)  'vgs))
       (* vdsat1 (((pd 2) exp) 'vdsat 'alpha )
  (((pd 1) alpha) 'vdsat)
  (((pd 1) vdsat) 'vgs))))
 (/ 1. (+ 1. delta)))
     ))))

       (cond ((< Vgs (+ Vfb Vbs))  (* W L Cox )) ;Accumulation
     ((< Vgs Vth) (f1 Vgs vbs vds))
     (else (f2 vgs vbs vds))))
     ))


(regions+ Basic-d/dvgs-Qg Vgs Vbs Vds Vth mu-zero W L))))
```

The result:

```
(define d/dvgs-qg
(lambda
 (vgs vbs vds vth mu-zero w l)
 (let
  ((basic-d/dvgs-qg
    (lambda
     (vgs vbs vds vth mu-zero w l)
     (define
      f1
      (lambda
       (vgs vbs vds)
       (*
        w
        l
        cox
        (let*
         ((q
           (+
            (- (* .5 gamma))
            (sqrt (+ (* gamma gamma .25) vgs (- vbs) (- vfb)))))
          (qvth
           (+
            (- (* .5 gamma))
            (sqrt (+ (* gamma gamma .25) vth (- vbs) (- vfb))))))
         (*
          (/ (* gamma (sqrt (+ phis (* -1 vbs)))) qvth)
          (/
```

```
        1/2
        (sqrt (+ (* .25 (expt gamma 2)) (* -1 vbs) (* -1 vfb) vgs))))))))
    (define
     f2
     (lambda
      (vgs vbs vds)
      (let*
       ((delta (/ gamma (* 2. (sqrt (- phis vbs)))))
        (vdsat (/ (- vgs vth) (+ 1. delta)))
        (alpha
         (if (and (<= vds vdsat) (> vdsat 0.)) (- 1. (/ vds vdsat)) 0.)))
       (*
        w
        l
        cox
        (if
         (> vdsat 0.)
         (+
          (*
           (+ delta (* 2/3 (/ (+ 1 alpha (* alpha alpha)) (+ 1 alpha))))
           (/ 1 (+ delta 1)))
          (*
           (/ (- vgs vth) (+ 1 delta))
           (*
            (*
             (/
              (+ (* 2/3 (expt alpha 2)) (* 4/3 alpha))
              (+ (expt alpha 2) (* 2 alpha) 1))
             (/ vds (expt vdsat 2)))
            (/ 1 (+ delta 1)))))
         (/ 1. (+ 1. delta)))))))
   (cond                              ((< vgs (+ vfb vbs)) (* w l cox))
    ((< vgs vth) (f1 vgs vbs vds)) (else (f2 vgs vbs vds))))))
(regions+ basic-d/dvgs-qg vgs vbs vds vth mu-zero w l))))
```

Acknowledgement

The cooperations of Project MAC (Mathematics and Computations) of the AI lab, MIT is gratefuly acknowledged: H. Abelson and G. J. Sussman have written the scheme circuit packages including the symbolic facilities; They also listened and commented on the material. C. P. Hansen and G. J. Rozas have written the scheme compiler and always extended help when needed.

Helpful discusions with Prof. A. Arbel, Dr. D. Lubzens, and Dr. L. Goldmintz of the Technion and Dr. A. Kolodny of Intel (Israel) are gratefuly acknowledged. Thanks are due also to Dr. Richard Zippel, Cornell University, for commenting on the manuscript.

This report describes research done at the Technion–Israel Institute of Technology and the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for this research is provided in part by the Technion Fund for the Promotion of Research, by the USA-Israel Bi-National Science Foundation, the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-92-J-4097 and by the National Science Foundation under grant number MIP-9001651.

References

[1] H. Abelson, A. Berlin, J. Katzenelson, W. McAllister, G. Rozas, and G.J. Sussman, "The Supercomputer Toolkit and its Applications," *Proc. Of the Fifth Jerusalem Conference on Information Technology*, Oct. 1990. Also available as AI Memo 1249.

[2] H. Abelson, A. Berlin, J. Katzenelson, W. McAllister, G. Rozas, G. J. Sussman, and J. Wisdom, "The Supercomputer Toolkit: A General Framework for Special–Purpose Computing," AI Memo 1329, Artificial Intelligence Laboratory, MIT, November 1991.

[3] Y.P. Tsividis, "Operation and Modeling of the MOS Transistor," McGraw Hill Book Co., 1987.

[4] D.E. Ward and R.W. Dutton, "A Charge-Oriented Model for the MOS Transistor Capacitances," IEEE Journal of Solid-State Circuits. Vol sc-13, October 1978.

[5] B.J. Sheu, D. L. Scharfetter, and P. K. Ko, "BSIM: Berekeley Short-Channel IGFET Model for MOS Transistor," IEEE Journal of Solid-State Circuits. Vol sc-22,no. 4 August 1987.

[6] B.J. Sheu, W-J Hsh and P. K. Ko, "An MOS Transistor Charge Model for VLSI Design," IEEE Trans. On Computer-Aided Design, Vol 7, No. 4, April 1988.

[7] P. Antognetti and G. Massobrio, "Semiconductor device modeling with SPICE," MCGraw-Hill, 1987.

[8] C.C. McAndrew, B. K. Bhattacharyya and O. Wing " A Single-Piece C_∞-Continuous MOSFET Model Including Subthreshold Conduction," IEEE Electronic Device Letters, Vol 12, No. 10, October 1991, pp. 565-567.

[9] D. A. Antoniadis, private communication.

[10] L.O. Chua And P.M. Lin "Computer Aided Analysis of Electronic Circuits," McGraw Hill Book Co, 1975.

[11] R.L. Geiger, P.E. Allen, and N.R. Strader, "VLSI Design Techniques for Analog and Digital Circuits," McGraw Hill Book Co., 1990.

[12] A. Arbel, private communication.

[13] J. Williams, "Analog Circuit Design," Butterworth-Heinemann, 1992.

[14] C. A. Desoer and J. Katzenelson, "Nonlinear RLC Networks," BSTJ, vol 44, pp. 161-198, January 1965.

[15] H. Abelson and J. Katzenelson, "Nonlinear N-terminals Networks," in preperation.

[16] R.S. Palais, "Natural Operations on Differential Forms," Trans. of the American Math. Soc., 92, 125-145, 1959.

[17] G. J. Minty, "Monotone (Nonlinear) Operators in Hilbert Space," Duke Math. J., 29 (1962), 341–346.

[18] C. L. Dolph and G. J. Minty, "On Nonlinear Intergral Equations of the Hammerstien Type," in "Nonlinear Integral Equations," P. M. Ansloe, Ed., University of Wisconsin Press, Madison, Wisconsin, 1964.

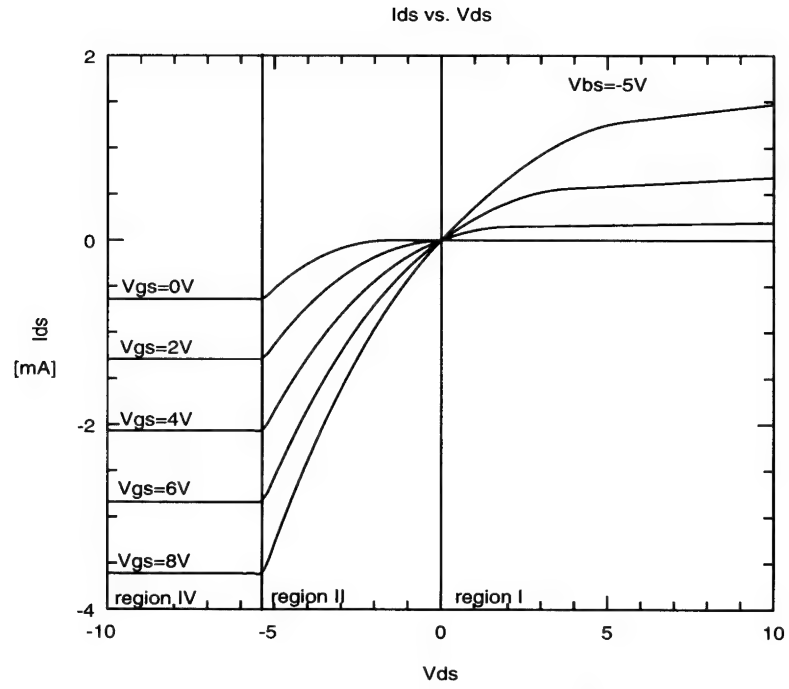

Figure 10: Ids for Vbs=-5V

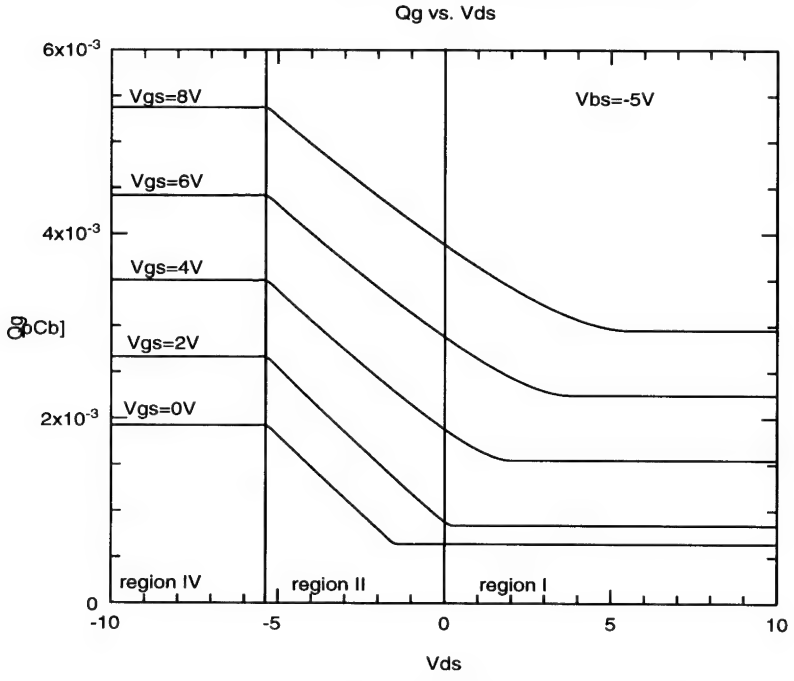

Figure 11: Qg for Vbs=-5V

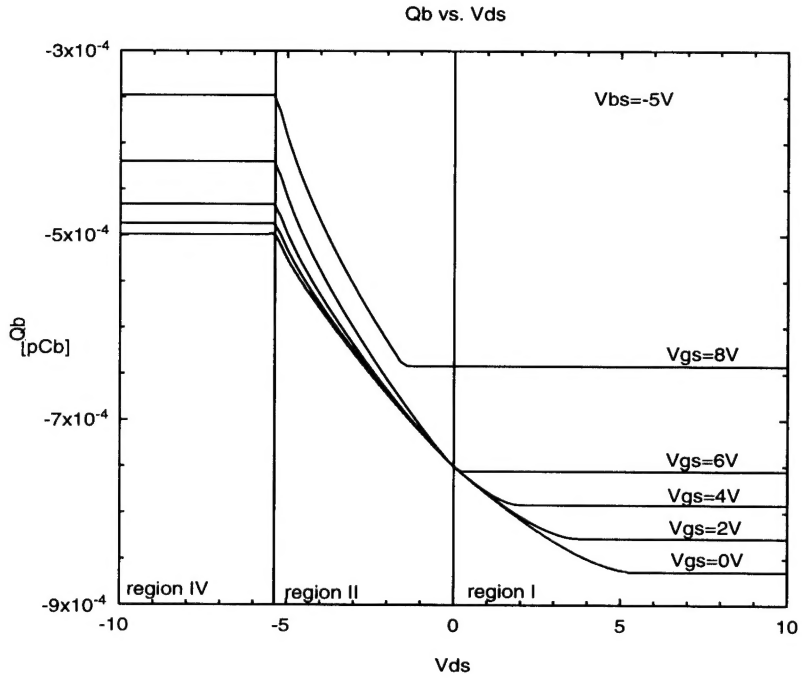

Figure 12: Qb for Vbs=-5V

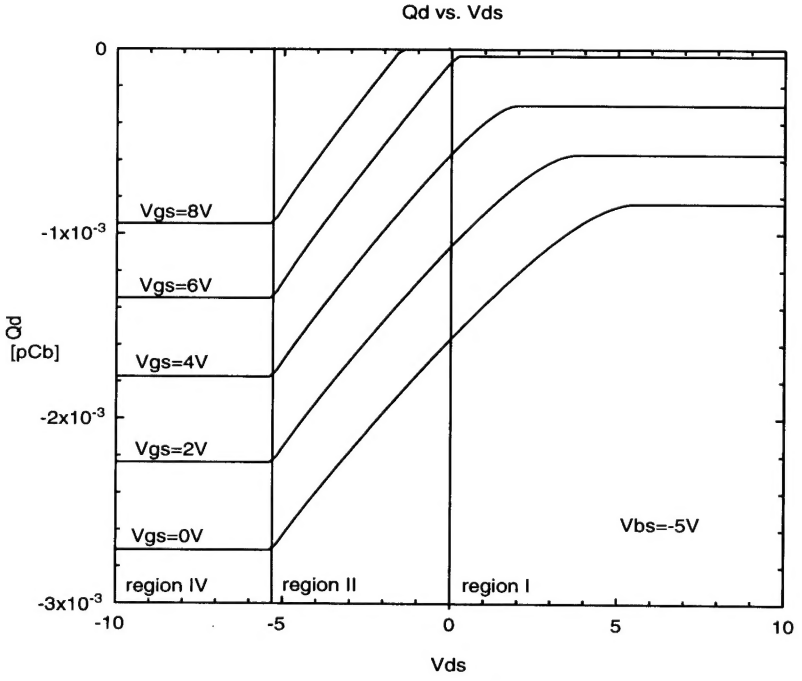

Figure 13: Qd for Vbs=-5V

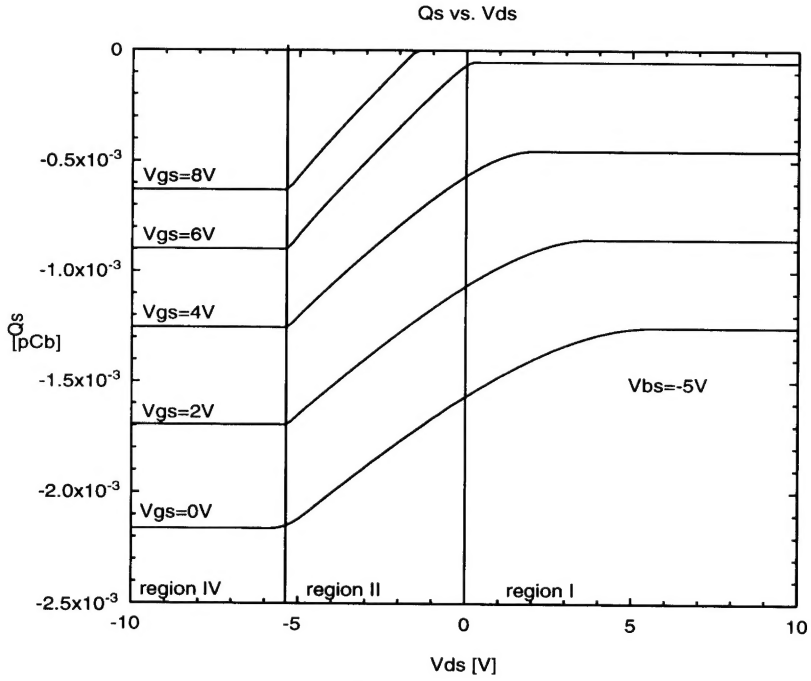

Figure 14: Qs for Vbs=-5V

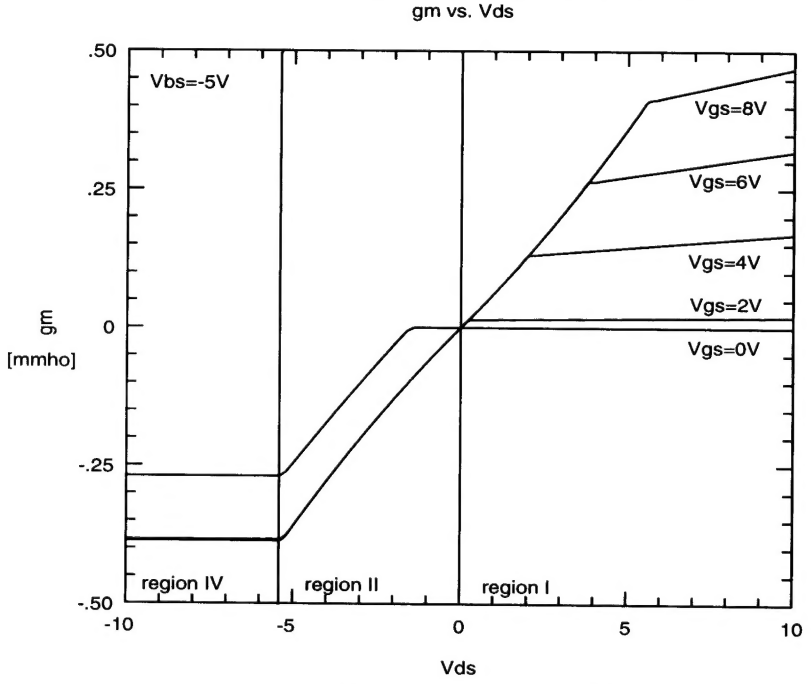

Figure 15: gm for Vbs=-5V

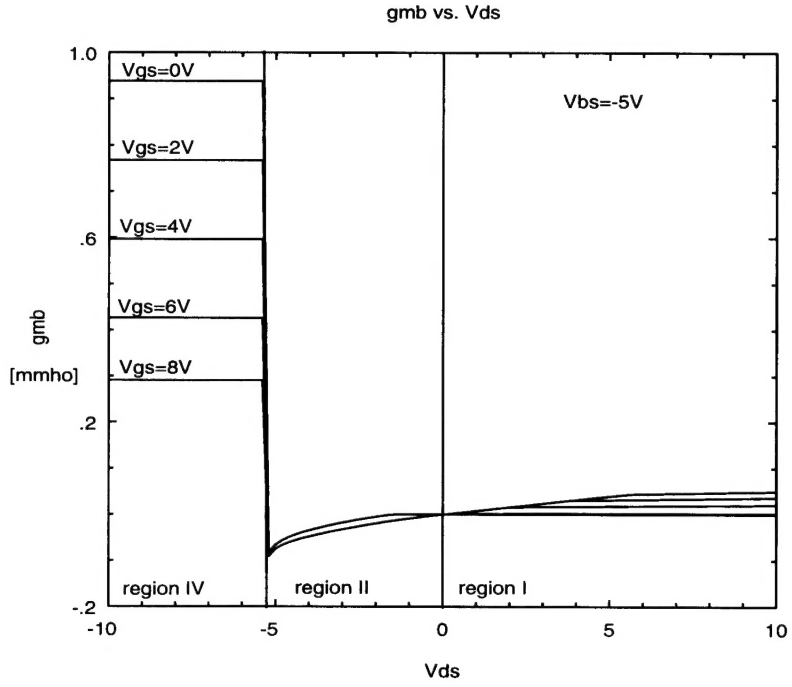

Figure 16: gmb for Vbs=-5V

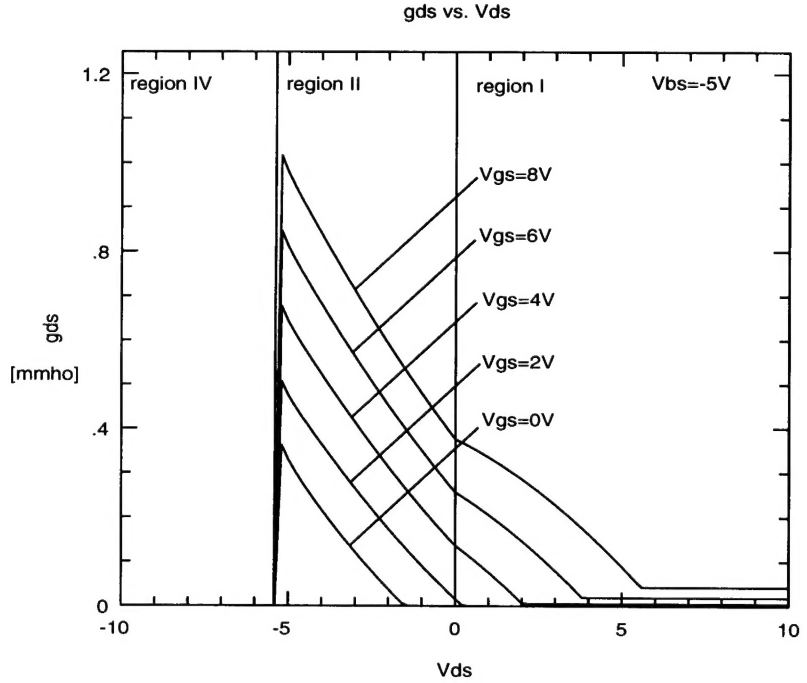

Figure 17: gds for Vbs=-5V

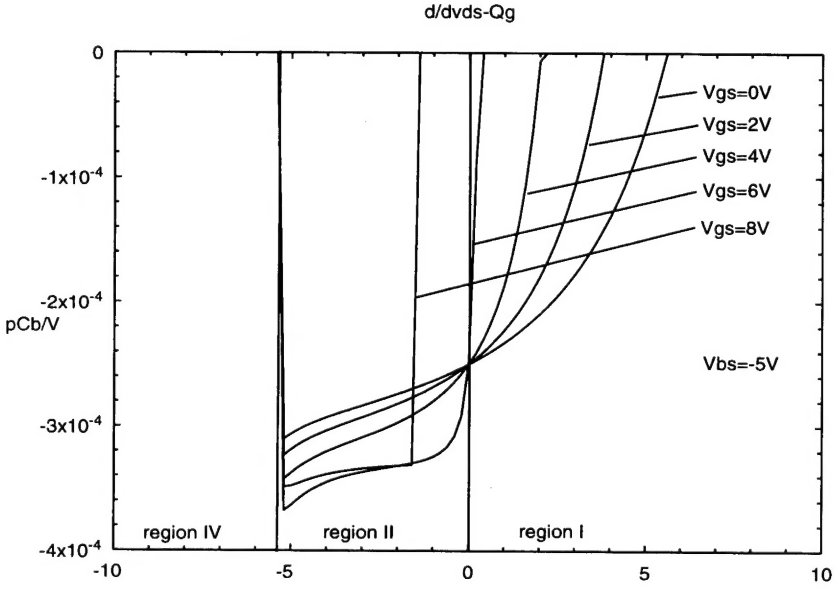

Figure 18: $\frac{dQg}{dV_{gs}}$ for Vbs=-5V